abiraterone+prednisone (AbP) improve overall survival (OS) in patients with mCRPC vs. mitoxantrone+P (MP) or P alone, respectively. Examination of patient and disease characteristics noted differences in the exposure to docetaxel and discontinuation of docetaxel due to progressive disease. The current study: 1) conducted systematic literature reviews of second-line treatment studies; 2) reviewed NICE and IQWIG submissions; 3) reviewed the initial and updated OS data from the TROPIC and COU-AA-301 studies; 4) interviewed clinical experts; and 5) performed a meta-analysis of two first-line (1L) mCRPC studies to inform the ITC on the OS for the two treatments and connect the network. Three comparisons were performed using hazard ratios (HRs) for the MP vs. P: 1.0 (clinical expert opinion), 0.97 (1L studies meta-analysis), and 0.90 (survival curve extraction). The Bucher ITC was used with a HR (CbzP vs. AbP)<1 favoring CbzP. **RESULTS:** Results based on updated OS data were consistent across methodologies, with HR(OS, clinical)=0.97 (95%CI: 0.78-1.21), HR(OS, meta-analysis)=0.95 (95%CI: 0.69-1.30) and HR(OS, extraction)=0.88 (95%CI: 0.63-1.21), but all HRs were not significantly different. This was observed in the docetaxel-resistant subgroup as well; with HR(OS, clinical) =0.95 (95%CI: 0.70-1.28), HR(OS, meta-analysis) =0.92 (95%CI: 0.63-1.34) and HR(OS, extraction)=0.85 (95%CI: 0.59-1.24). These are different from the results presented in the IQWIG submission assuming that MP has the same effect as P alone based on initial OS data. **CONCLUSIONS:** Differences in results highlight the dependency of ITCs on efficacy assumptions. Lack of equivalence in disease, comparators or patient characteristics contribute to uncertainty regarding conclusions, which further emphasizes the fact that randomized prospective clinical trials are best suited to fully evaluate the efficacy and safety of cancer treatments.

RESEARCH ON METHODS - Patient-Reported Outcomes Studies

PRM87
CROSS-CULTURAL VALIDATION OF THE HUNTINGTON QUALITY OF LIFE INSTRUMENT IN SPAIN

Khemiri A[1], Clay E[2], Trigo P[3], Yebenes J[3], Toumi M[4]
[1]*Creativ-Ceutical, Tunis, Tunisia,* [2]*Creativ-Ceutical, Paris, France,* [3]*Hospital Ramón y Cajal, Madrid, Spain,* [4]*University Claude Bernard Lyon 1, Lyon, France*

OBJECTIVES: The Huntington Quality of Life Instrument (H-QoL-I) is the first self-reported, disease-specific instrument developed to assess the health-related quality of life (HRQoL) of patients with Huntington's disease (HD). It was originally developed and validated for France and Italy, and later also validated for Germany, Poland and the USA. This study aimed to validate the Spanish version of H-QoL-I cross-culturally. **METHODS:** The original questionnaire included three subscales assessing motor functioning, psychology and socializing (11 items). The instrument was translated forwards and backwards by native speakers. A survey was conducted with 59 patients. Face validity was tested through item completion and overall understanding. Internal validity was tested, assessing internal consistency, correlation matrix using item/dimension correlation and factorial structure. External validation was performed versus motor symptoms, behavioral symptoms, independence, and the EuroQoL 5D (EQ-5D). Differential item functioning (DIF) analyses were performed versus data from the Italian and French versions, using Zumbo criteria. **RESULTS:** Item response rates ranged from 87% to 97%. A floor effect was found for three items. Results showed that the scale had a good reliability (Cronbach's alpha coefficients > 0.75). Factor analyses demonstrated satisfactory construct validity. Item internal consistency (IIC) and item discriminant validity criteria were met for most items (i.e. IIC was > 0.40, and correlations between items and their respective rest-scores in one dimension were all greater than correlations with another dimension). The external validity was supported by correlation of the different dimensions with the related clinical symptoms and related generic QoL dimensions. The correlation between total H-QoL-I score and EQ-5D index score was 0.78. No DIF was detected. **CONCLUSIONS:** These data support the cross-cultural validity of the H-QoL-I to assess the health status of patients with HD and integrate the patient perspective for Spain. A limitation of this study is the small sample size.

PRM88
TRANSLATION AND VALIDATION OF OSTEOPOROSIS HEALTH BELIEF SCALE INTO MALAYSIAN VERSION AMONG TYPE 2 DIABETICS PATIENTS

Abdulameer SA[1], Syed Sulaiman SA[1], Hassali MA[1], Subramaniam K[2], Sahib MN[1]
[1]*Universiti Sains Malaysia, Minden, Penang, Malaysia,* [2]*Penang General Hospital, George Town, Penang, Malaysia*

OBJECTIVES: To translate and examine the psychometric properties of the Malaysian version of the Osteoporosis Health Belief Scale (OHBS-M) among type 2 diabetes patients (T2DM) and to determine the best cut-off value for OHBS-M with optimum sensitivity and specificity. **METHODS:** A standard "forward–backward" procedure was used to translate OHBS into Malay language. It was later validated on a convenience sample of 250 T2DM outpatients between May and July 2011. The psychometric assessment of this study was including validity (face validity, content validity ratio, and construct validity) and reliability (internal consistency and test-retest). Sensitivity and specificity of OHBS-M were calculated using receiver operating characteristic (ROC) curve analysis in comparison with the proxy gold standard (quantitative ultrasound scan). **RESULTS:** The mean ±SD of OHBS-M scores was 158.31± 20.80. The Fleiss' kappa, content validity ratio range and content validity index were 0.99, 0.75 to 1 and 0.886, respectively. Seven factors of the OHBS-M were identified using exploratory factor analysis and were confirmed through confirmatory factor analysis. Internal consistency and test–retest reliability value were 0.89 and 0.61, respectively. The cut-off value of the OHBS-M was 169 with a sensitivity of 77.4% (95% CI 0.68- 0.84) and a specificity of 78.2% (95% CI 0.69-0.85) to identify osteoporosis/osteopenia patients. The positive and negative predictive values were 78% (95% CI 0.68 - 0.85) and 77.6% (95% CI 0.68-0.84), respectively. The area under the curve (AUC) for the OHBS-M was 0.877(95% CI 0.82-0.92). According to QUS measurements, 20.4% were considered osteoporotic, while, 57.6 % osteopenic and 22 % normal. **CONCLUSIONS:** The findings of this study suggest that the OHBS-M instrument is valid and reliable tool to be used in Malaysian clinical setting.

PRM89
STANDARDIZING THE METRIC AND INCREASING THE EFFICIENCY OF PHYSICAL FUNCTIONING OUTCOMES MEASUREMENT

Ware J[1], Guyer R[2], Harrington M[2], Boulanger R[2]
[1]*University of Massachusetts Medical School, Worcester, MA, USA,* [2]*John Ware Research Group, Worcester, MA, USA*

OBJECTIVES: Item response theory (IRT) modeling evaluated the metric underlying generic physical function (PF) surveys including SF-36, PROMIS and new PF categorical rating items and tested whether scores could be estimated more efficiently while maintaining forward-backward comparability. **METHODS:** Generalized Partial Credit Model (GPCM) estimates of parameters for MOS SF-36 PF-10, PROMIS 6-item PF and new (easy-hard) PF categorical rating items and model fit were tested in a probability sample representing the general US population (N = 625). Analyses included: (a) fit of GPCM for 35-item bank; (b) item utilization in computerized adaptive tests (CAT), (c) % at ceiling and floor; (d) % for whom reliability > 0.90 (reliable range); (e) equivalence of mean norm–based scores (mean=50, SD=10) for all measures across mild, moderate and severe chronically-ill groups, and (f) validity in predicting physical and emotional health general summary measures at a 9-month follow-up. **RESULTS:** The GPCM fit the data and item parameter estimates agreed very well with those previously reported for MOS and PROMIS PF items. In tests of discriminant validity, group means differed substantially across severity groups (RV = 0.81 to 1.00) and score equivalence across methods within each group was confirmed (all differences < 1 point). RV's for standardized PF scores estimated from new E-H items were equivalent to PF-10 and PROMIS PF estimations. Predictive validity was equivalent and substantial (across methods) for physical and significant, but lower, for emotional outcomes at 9 months, as hypothesized. The most efficient (reduced respondent burden, comparable or improved reliability and validity) measure was a new 6-item PF using E-H items and an improved adaptive survey logic. **CONCLUSIONS:** Findings support the standardization of the metric underlying PF measures and extend choice of methods to include more efficient categorical rating scales that maintain forward-backward score comparability. Improved adaptive survey logic reduces respondent burden and increases the reliable range for estimates of scores for familiar legacy measures. This approach warrants application to other generic health domains and tests of translated items and standardized parameters across countries and languages.

PRM90
OUTCOME DIFFERENCES IN ALGORITHMS USED FOR INDIRECT MAPPING OF UTILITY VALUES FROM HAQ-DI: AN ASSESSMENT BASED ON PHASE 2A CLINICAL TRIAL DATA IN PATIENTS WITH RHEUMATOID ARTHRITIS AFTER TREATMENT WITH NNC0109-0012 (ANTI-IL-20 MAB)

Strandberg-Larsen M[1], Hansen BB[1], Göthberg M[1], Valencia X[2]
[1]*Novo Nordisk A/S, Søborg, Denmark,* [2]*Novo Nordisk Inc, Princeton, NJ, USA*

OBJECTIVES: A utility value is a preference-based measure for a person's health-related quality of life at a given point in time. In clinical trials the standard is to measure utilities based on e.g. EQ-5D or HUI-3. However, in rheumatoid arthritis (RA) there has been a tradition of estimating utility values indirectly from HAQ-DI. This study used data from a phase 2a trial in RA patients after treatment with NNC0109-0012 to assess outcome differences applying various indirect mapping algorithms. **METHODS:** At least five different indirect mapping algorithms (UV1-5) that translates HAQ-DI to EQ-5D or HUI3 utility values have been published. These were applied to a phase 2a, multicentre, randomised, double-blind, multiple-dose, placebo-controlled, parallel group trial investigating the clinical efficacy of NNC0109-0012 in RA patients with active disease (Results reported elsewhere). Physical function was a secondary objective measured by the change in HAQ-DI from baseline to week 12. The analysis was performed with an ANOVA with treatment as fixed factor. The following five algorithms were used: UV1(EQ-5D)=0.9567–0.309*HAQ-DI, Min;Max=(0.03;0.96); UV2(HUI3)=0.76–0.28*HAQ-DI+0.05*(if Female), Min;Max=(-0.08;0.81); UV3(HUI3)=0.76–0.28*HAQ-DI+0.05*(if Female)+0.001*Age, Min;Max=(-0.06;0.91); UV4(HUI3)=0.74–0.17*HAQ-DI, Min;Max=(0.23;0.74); and UV5(HUI3)=0.9527–0.2018*HAQ-DI, Min;Max=(0.35;0.95). **RESULTS:** After 12 weeks the mean utility improvement within the active-group across algorithms was 0.11 (range: 0.08-0.14) and 0.03 (range: 0.02-0.05) in the placebo-group. When comparing utility improvements between active and placebo group across algorithms the difference was most pronounced when using UV3 = 0.10 (range: 0.06-0.10) taking HAQ-DI, sex and age into account. **CONCLUSIONS:** Choice of mapping algorithm for conversion of HAQ-DI into utilities impacts the outcome in term of utility improvements, although the differences are small. Future clinical trials using direct assessment of utilities will substantiate the potential benefits of NNC0109-0012 for patients suffering from RA. Direct elicitations can also be used to shed additional light on the validity of available indirect mapping algorithms.

PRM91
HOW DO PATIENTS WITH DIFFERENT CONDITIONS DESCRIBE THEIR PAIN?

Martin ML[1], Scanlon M[2], McCarrier KP[1], Wolfe M[2], Bushnell DM[2]
[1]*Health Research Associates, Inc., Seattle, WA, USA,* [2]*Health Research Associates, Inc., Mountlake Terrace, WA, USA*

OBJECTIVES: To identify descriptors that patients who have different physical conditions use to describe the quality and severity of their pain, and to examine com-